Extreme Lp-quantile kernel regression

Stéphane Girard, Gilles Stupfler, Antoine Usseglio-Carleve

Extreme L^p−quantile kernel regression

Stéphane Girard, Gilles Stupfler and Antoine Usseglio-Carleve

Abstract Quantiles are recognized tools for risk management and can be seen as minimizers of an L^1−loss function, but do not define coherent risk measures in general. Expectiles, meanwhile, are minimizers of an L^2−loss function and define coherent risk measures; they have started to be considered as good alternatives to quantiles in insurance and finance. Quantiles and expectiles belong to the wider family of L^p−quantiles. We propose here to construct kernel estimators of extreme conditional L^p−quantiles. We study their asymptotic properties in the context of conditional heavy-tailed distributions and we show through a simulation study that taking $p \in (1, 2)$ may allow to recover extreme conditional quantiles and expectiles accurately. Our estimators are also showcased on a real insurance data set.

1 Introduction

The quantile, also called Value-at-Risk in actuarial and financial areas, is a widespread tool for risk measurement, due to its simplicity and interpretability: if Y is a random variable with a cumulative distribution function F, the quantile at level $\alpha \in (0, 1)$ is defined as $q(\alpha) = \inf \{y \in \mathbb{R} | F(y) \geq \alpha\}$. As pointed out in [19], quantiles may also be seen as a solution of the following minimization problem:

$$q(\alpha) = \underset{t \in \mathbb{R}}{\arg \min} \, \mathbb{E} \left[\rho_\alpha^{(1)}(Y - t) - \rho_\alpha^{(1)}(Y) \right], \tag{1}$$

where $\rho_\alpha^{(1)}(y) = |\alpha - \mathbb{1}_{\{y \leq 0\}}||y|$ is the quantile check function. However, the quantile is not subadditive in general and so is not a coherent risk measure in the sense of [1]. An alternative risk measure gaining popularity is the expectile, introduced in [20]. This is the solution of (1), with the new loss function $\rho_\alpha^{(2)}(y) = |\alpha - \mathbb{1}_{\{y \leq 0\}}|y^2$ in place of $\rho_\alpha^{(1)}$. Expectiles larger than the mean are coherent risk measures, and have started to be used in actuarial and financial practice (see for instance [3]). A pioneering paper for the estimation of extreme expectiles in heavy-tailed settings is [7].

Stéphane Girard
Univ. Grenoble-Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France, e-mail: stephane.girard@inria.fr

Gilles Stupfler
Univ Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes, France, e-mail: gilles.stupfler@ensai.fr

Antoine Usseglio-Carleve
Univ. Grenoble-Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France, e-mail: antoine.usseglio-carleve@inria.fr

Quantiles and expectiles may be generalized by considering the family of L^p−quantiles. Introduced in [4], this class of risk measures is defined, for all $p \geq 1$, by:

$$q^{(p)}(\alpha) = \arg\min_{t \in \mathbb{R}} \mathbb{E}\left[\rho_\alpha^{(p)}(Y - t) - \rho_\alpha^{(p)}(Y)\right], \tag{2}$$

where $\rho_\alpha^{(p)}(y) = |\alpha - \mathbb{1}_{\{y \leq 0\}}||y|^p$ is the L^p−quantile loss function; the case $p = 1$ leads to the quantile and $p = 2$ gives the expectile. Note that, for $p > 1$, using the formulation (2) and through the subtraction of the (at first sight unimportant) term $\rho_\alpha^{(p)}(Y)$, it is a straightforward consequence of the mean value theorem applied to the function $\rho_\alpha^{(p)}$ that the L^p−quantile $q^{(p)}(\alpha)$ is well-defined as soon as $\mathbb{E}(|Y|^{p-1}) < \infty$. While the expectile is the only coherent L^p−quantile (see [2]), [8] showed that for extreme levels of quantiles or expectiles ($\alpha \to 1$), it may be better to estimate L^p−quantiles first (where typically p is between 1 and 2) and exploit an asymptotic proportionality relationship to estimate quantiles or expectiles. An overview of the potential applications of this kind of statistical assessment of extreme risk may for instance be found in [14].

The contribution of this work is to propose a methodology to estimate extreme L^p−quantiles of $Y|\mathbf{X} = \mathbf{x}$, where the random covariate vector $\mathbf{X} \in \mathbb{R}^d$ is recorded alongside Y. In this context, the case $p = 1$ (quantile) has been considered in [6] and [5], and the case $p = 2$ (expectile) has recently been studied in [17]. For the general case $p \geq 1$, only [22] proposes an estimation procedure under the strong assumption that the vector $(\mathbf{X}, Y)$ is elliptically distributed. The present paper avoids this modeling assumption by constructing a kernel estimator.

The paper is organized as follows. Section 2 introduces an estimator of conditional L^p−quantiles. Section 3 gives the asymptotic properties of the estimator previously introduced, at extreme levels. Finally, Section 4 proposes a simulation study in order to assess the accuracy of our estimator which is then showcased on a real insurance data set in Section 5. Proofs are postponed to the Appendix.

2 L^p−quantile kernel regression

Let $(\mathbf{X}_i, Y_i)$, $i = 1, \ldots, n$ be independent realizations of a random vector $(\mathbf{X}, Y) \in \mathbb{R}^d \times \mathbb{R}$. For the sake of simplicity we assume that $Y \geq 0$ with probability 1. We denote by g the density function of $\mathbf{X}$ and let, in the sequel, $\mathbf{x}$ be a fixed point in $\mathbb{R}^d$ such that $g(\mathbf{x}) > 0$. We denote by $\bar{F}^{(1)}(y|\mathbf{x}) = \mathbb{P}(Y > y|\mathbf{X} = \mathbf{x})$ the conditional survival function of Y given $\mathbf{X} = \mathbf{x}$ and assume that this survival function is continuous and regularly varying with index $-1/\gamma(\mathbf{x})$:

$$\forall t > 0, \ \lim_{y \to \infty} \frac{\bar{F}^{(1)}(ty|\mathbf{x})}{\bar{F}^{(1)}(y|\mathbf{x})} = t^{-1/\gamma(\mathbf{x})}. \tag{3}$$

Such a distribution belongs to the Fréchet maximum domain of attraction [11]. Note that for any $k < 1/\gamma(\mathbf{x})$, $\mathbb{E}\left[Y^k|\mathbf{X} = \mathbf{x}\right] < \infty$. Since the definition of L^p−quantiles in (2) requires $\mathbb{E}\left[|Y|^{p-1}|\mathbf{X} = \mathbf{x}\right] < \infty$, our minimal assumption will be that $p - 1 <$

$1/\gamma(\mathbf{x})$. From Equation (2), L^p–quantiles of level $\alpha \in (0, 1)$ of Y given $\mathbf{X} = \mathbf{x}$ may also be seen as the solution of the following equation:

$$\frac{\mathbb{E}\left[|Y - y|^{p-1}\, \mathbb{1}_{\{Y>y\}}|\mathbf{X} = \mathbf{x}\right]}{\mathbb{E}\left[|Y - y|^{p-1}\,|\mathbf{X} = \mathbf{x}\right]} = 1 - \alpha.$$

In other terms, as noticed in [18], (conditional) L^p–quantiles can be equivalently defined as quantiles

$$q^{(p)}(\alpha|\mathbf{x}) = \inf\left\{y \in \mathbb{R} \mid \bar{F}^{(p)}(y|\mathbf{x}) \le 1 - \alpha\right\}$$

of the distribution associated with the survival function

$$\bar{F}^{(p)}(y|\mathbf{x}) = \frac{\varphi^{(p-1)}(y|\mathbf{x})}{m^{(p-1)}(y|\mathbf{x})},$$

where, for all $k \ge 0$,

$$m^{(k)}(y|\mathbf{x}) = \mathbb{E}\left[|Y - y|^k\,|\mathbf{X} = \mathbf{x}\right] g(\mathbf{x})$$

$$\text{and } \varphi^{(k)}(y|\mathbf{x}) = \mathbb{E}\left[|Y - y|^k\, \mathbb{1}_{\{Y>y\}}|\mathbf{X} = \mathbf{x}\right] g(\mathbf{x}).$$

Obviously, if $p = 1$, we get the survival function introduced above. The case $p = 2$ leads to the function introduced in [18] and used in [17]. To estimate $\bar{F}^{(p)}(y|\mathbf{x})$, we let K be a probability density function on $\mathbb{R}^d$ and we introduce the kernel estimators

$$\hat{m}_n^{(k)}(y|\mathbf{x}) = \frac{\sum_{i=1}^{n} |Y_i - y|^k\, K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)}{nh_n^d}, \quad \hat{\varphi}_n^{(k)}(y|\mathbf{x}) = \frac{\sum_{i=1}^{n} |Y_i - y|^k\, K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right) \mathbb{1}_{\{Y_i>y\}}}{nh_n^d}.$$

Note that $\hat{m}_n^{(0)}(0|\mathbf{x})$ is the kernel density estimator of $g(\mathbf{x})$, and $\hat{m}_n^{(1)}(0|\mathbf{x})/\hat{m}_n^{(0)}(0|\mathbf{x})$ is the standard kernel regression estimator (since the Y_i are nonnegative). The kernel estimators of $\bar{F}^{(p)}(y|\mathbf{x})$ and $q^{(p)}(\alpha|\mathbf{x})$ are then easily deduced:

$$\hat{\bar{F}}_n^{(p)}(y|\mathbf{x}) = \frac{\hat{\varphi}_n^{(p-1)}(y|\mathbf{x})}{\hat{m}_n^{(p-1)}(y|\mathbf{x})}, \quad \hat{q}_n^{(p)}(\alpha|\mathbf{x}) = \inf\left\{y \in \mathbb{R} \mid \hat{\bar{F}}_n^{(p)}(y|\mathbf{x}) \le 1 - \alpha\right\}. \quad (4)$$

The case $p = 1$ gives the kernel quantile estimator introduced in [5], while $p = 2$ leads to the conditional expectile estimator of [17]. We study here the asymptotic properties of $\hat{q}_n^{(p)}(\alpha|\mathbf{x})$ for an arbitrary $p \ge 1$, when $\alpha = \alpha_n \to 1$.

3 Main results

We first make a standard assumption on the kernel. We fix a norm $||\cdot||$ on $\mathbb{R}^d$.

$(\mathcal{K})$ The density function K is bounded and its support S is contained in the unit ball.

To be able to analyze extreme conditional L^p-quantiles in a reasonably simple way, we make a standard second-order regular variation assumption (for a survey of those conditions, see Section 2 in [11]).

C_2 $(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ There exist $\gamma(\mathbf{x}) > 0$, $\rho(\mathbf{x}) \leq 0$ and a positive or negative function $A(\cdot|\mathbf{x})$ converging to 0 such that:

$$\forall t > 0, \ \lim_{y \to \infty} \frac{1}{A(y|\mathbf{x})} \left(\frac{q^{(1)}(1 - 1/(ty)|\mathbf{x})}{q^{(1)}(1 - 1/y|\mathbf{x})} - t^{\gamma(\mathbf{x})} \right) = \begin{cases} t^{\gamma(\mathbf{x})} \dfrac{t^{\rho(\mathbf{x})} - 1}{\rho(\mathbf{x})} & \text{if } \rho(\mathbf{x}) < 0, \\ t^{\gamma(\mathbf{x})} \log(t) & \text{if } \rho(\mathbf{x}) = 0. \end{cases}$$

Our last assumption is a local Lipschitz condition which may be found for instance in [5, 12]. We denote by $B(\mathbf{x}, r)$ the ball with center $\mathbf{x}$ and radius r.

$(\mathcal{L})$ We have $g(\mathbf{x}) > 0$ and there exist $c, r > 0$ such that

$$\forall \mathbf{x}' \in B(\mathbf{x}, r), \ |g(\mathbf{x}) - g(\mathbf{x}')| \leq c||\mathbf{x} - \mathbf{x}'||.$$

To be able to control the local oscillations of $(\mathbf{x}, y) \mapsto \bar{F}^{(1)}(y|\mathbf{x})$, we let, for any nonnegative $y_n \to \infty$,

$$\omega_{h_n}^{(1)}(y_n|\mathbf{x}) = \sup_{\mathbf{x}' \in B(\mathbf{x}, h_n)} \sup_{z \geq y_n} \frac{1}{\log(z)} \left| \log \left(\frac{\bar{F}^{(1)}(z|\mathbf{x}')}{\bar{F}^{(1)}(z|\mathbf{x})} \right) \right|,$$

$$\omega_{h_n}^{(2)}(y_n|\mathbf{x}) = \sup_{\mathbf{x}' \in B(\mathbf{x}, h_n)} \sup_{0 < y \leq y_n} |\bar{F}^{(1)}(y|\mathbf{x}') - \bar{F}^{(1)}(y|\mathbf{x})|,$$

$$\text{and } \omega_{h_n}^{(3)}(y_n|\mathbf{x}) = \sup_{\mathbf{x}' \in B(\mathbf{x}, h_n)} \sup_{\lambda \geq 1} \sup_{b_n, b_n' \to 0} \left| \frac{\bar{F}^{(1)}(\lambda y_n(1 + b_n)|\mathbf{x}')}{\bar{F}^{(1)}(\lambda y_n(1 + b_n')|\mathbf{x}')} - 1 \right|.$$

The quantity $\omega_{h_n}^{(1)}(y_n|\mathbf{x})$, discussed for instance in [17], controls the oscillation of the conditional survival function with respect to $\mathbf{x}$ in its right tail, while $\omega_{h_n}^{(2)}(y_n|\mathbf{x})$ and $\omega_{h_n}^{(3)}(y_n|\mathbf{x})$ are introduced to be able to deal with the case $p \notin \{1, 2\}$ specifically. Let us highlight that $\omega_{h_n}^{(3)}(y_n|\mathbf{x})$ is again geared towards controlling an oscillation of the right tail of the conditional distribution; however, $\omega_{h_n}^{(2)}(y_n|\mathbf{x})$ focuses on the oscillation of the center of the conditional distribution with respect to $\mathbf{x}$. For $p > 1$, the introduction of a quantity such as $\omega_{h_n}^{(2)}(y_n|\mathbf{x})$ is in some sense natural, since we will have to deal with the local oscillation of the conditional moment $m^{(p-1)}(y|\mathbf{x})$, appearing in the denominator of $\bar{F}^{(p)}(y|\mathbf{x})$, and this conditional moment indeed depends on the whole of the conditional distribution rather than merely on its right tail. Typically $\omega_{h_n}^{(1)}(y_n|\mathbf{x}) = O(h_n)$, $\omega_{h_n}^{(2)}(y_n|\mathbf{x}) = O(h_n)$ and $\omega_{h_n}^{(3)}(y_n|\mathbf{x}) = o(1)$ under reasonable assumptions; we give examples below.

Remark 1 Assume that $Y|\mathbf{X} = \mathbf{x}$ has a Pareto distribution with tail index $\gamma(\mathbf{x}) > 0$:

$$\forall y \geq 1, \ \bar{F}^{(1)}(y|\mathbf{x}) = y^{-1/\gamma(\mathbf{x})}.$$

If γ is locally Lipschitz continuous, we clearly have $\omega_{h_n}^{(1)}(y_n|\mathbf{x}) = O(h_n)$. Further-more, for any $y \geq 1$, the mean value theorem yields

$$|\bar{F}^{(1)}(y|\mathbf{x}') - \bar{F}^{(1)}(y|\mathbf{x})| \leq \left|\frac{1}{\gamma(\mathbf{x}')} - \frac{1}{\gamma(\mathbf{x})}\right| \times y^{-1/[\gamma(\mathbf{x})\vee\gamma(\mathbf{x}')]} \log y.$$

(Here and below $\vee$ denotes the maximum operator.) Under this same local Lipschitz assumption, one then finds $\omega_{h_n}^{(2)}(y_n|\mathbf{x}) = O(h_n)$ as well. Finally, for any $y, y' > 1$,

$$\left|\frac{\bar{F}^{(1)}(y'|\mathbf{x}')}{\bar{F}^{(1)}(y|\mathbf{x}')} - 1\right| = \left|\left(\frac{y}{y'}\right)^{1/\gamma(\mathbf{x}')} - 1\right| \leq \frac{|y - y'|}{y'} \times \frac{1 + (y/y')^{1/\gamma(\mathbf{x}')-1}}{\gamma(\mathbf{x}')}$$

by the mean value theorem again. This inequality yields $\omega_{h_n}^{(3)}(y_n|\mathbf{x}) = o(1)$.

The same arguments, and asymptotic bounds on $\omega_{h_n}^{(1)}(y_n|\mathbf{x}), \omega_{h_n}^{(2)}(y_n|\mathbf{x})$ and $\omega_{h_n}^{(3)}(y_n|\mathbf{x})$, apply to the conditional Fréchet model

$$\forall y > 0, \ \bar{F}^{(1)}(y|\mathbf{x}) = 1 - \exp(-y^{-1/\gamma(\mathbf{x})}).$$

Analogous results are easily obtained for the conditional Burr model

$$\forall y > 0, \ \bar{F}^{(1)}(y|\mathbf{x}) = (1 + y^{-\rho(\mathbf{x})/\gamma(\mathbf{x})})^{1/\rho(\mathbf{x})}$$

when $\rho < 0$ is assumed to be locally Lipschitz continuous, and the conditional mixture Pareto model

$$\forall y \geq 1, \ \bar{F}^{(1)}(y|\mathbf{x}) = y^{-1/\gamma(\mathbf{x})} \left[c(\mathbf{x}) + (1 - c(\mathbf{x}))y^{\rho(\mathbf{x})/\gamma(\mathbf{x})}\right]$$

when $\rho < 0$ and $c \in (0, 1)$ are assumed to be locally Lipschitz continuous. □

3.1 Intermediate L^p−quantile regression

In this paragraph, we assume that $\sigma_n^{-2} = nh_n^d(1 - \alpha_n) \to \infty$. Such an assumption means that the L^p−quantile level α_n tends to 1 slowly (by extreme value standards), hence the denominations *intermediate sequence* and *intermediate L^p−quantiles*. This assumption is widespread in the literature of risk measure regression: see, among others, [5, 6, 12, 17]. Throughout, we let $\|K\|_2^2 = \int_S K(\mathbf{u})^2 d\mathbf{u}$ be the squared L^2−norm of K, $\Psi(\cdot)$ denote the digamma function and $IB(t, x, y) = \int_0^t u^{x-1}(1 - u)^{y-1}du$ be the incomplete Beta function. Note that $IB(1, x, y) = B(x, y)$ is the standard Beta function.

We now give our first result on the joint asymptotic normality of a finite number J of empirical conditional quantiles with an empirical conditional L^p-quantile ($p > 1$).

Theorem 1 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and C_2 $(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ hold. Let $\alpha_n \to 1$, $h_n \to 0$ and $a_n = 1 - \tau(1 - \alpha_n)(1 + o(1))$, where $\tau > 0$. Assume further that

$\sigma_n^{-2} = nh_n^d(1 - \alpha_n) \to \infty$, $nh_n^{d+2}(1 - \alpha_n) \to 0$, $\sigma_n^{-1}A\left((1 - \alpha_n)^{-1}|\mathbf{x}\right) = O(1)$,
$\omega_{h_n}^{(3)}(q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}) \to 0$ and there exists $\delta \in (0, 1)$ such that

$$\sigma_n^{-1}\omega_{h_n}^{(1)}((1 - \delta)(\theta \wedge 1)q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}) \log(1 - \alpha_n) \to 0, \tag{5}$$

where $\theta = \left(\tau\gamma(\mathbf{x})/B\left(p, \gamma(\mathbf{x})^{-1} - p + 1\right)\right)^{-\gamma(\mathbf{x})}$. Let further $\alpha_{n,j} = 1 - \tau_j(1 - \alpha_n)$, for $0 < \tau_1 < \tau_2 < \ldots < \tau_J \leq 1$ such that

$$\sigma_n^{-1}\omega_{h_n}^{(2)}((1 + \delta)(\theta \vee \tau_1^{-\gamma(\mathbf{x})})q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}) \to 0. \tag{6}$$

Then, for all $p \in (1, \gamma(\mathbf{x})^{-1}/2 + 1)$, one has

$$\sigma_n^{-1}\left\{\left(\frac{\hat{q}_n^{(1)}(\alpha_{n,j}|\mathbf{x})}{q^{(1)}(\alpha_{n,j}|\mathbf{x})} - 1\right)_{1\leq j\leq J}, \left(\frac{\hat{q}_n^{(p)}(a_n|\mathbf{x})}{q^{(p)}(a_n|\mathbf{x})} - 1\right)\right\} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}_{J+1}, \frac{||K||_2^2}{g(\mathbf{x})}\gamma(\mathbf{x})^2\mathbf{\Sigma}(\mathbf{x})\right), \tag{7}$$

where $\mathbf{\Sigma}(\mathbf{x})$ is the symmetric matrix having entries

$$\begin{cases} \Sigma_{j,\ell}(\mathbf{x}) = & \left(\tau_j \vee \tau_\ell\right)^{-1} \\ \Sigma_{j,J+1}(\mathbf{x}) = & \tau_j^{-1}\left[\gamma(\mathbf{x})\dfrac{(p-1)IB\left(\left(1\vee\frac{\tau_j^{-\gamma(\mathbf{x})}}{\theta}\right)^{-1},\gamma(\mathbf{x})^{-1}-p+1,p-1\right)}{B(p,\gamma(\mathbf{x})^{-1}-p+1)} + \left(\left(1 \vee \frac{\tau_j^{-\gamma(\mathbf{x})}}{\theta}\right) - 1\right)^{p-1}\right] \\ \Sigma_{J+1,J+1}(\mathbf{x}) = & \dfrac{B\left(2p-1,\gamma(\mathbf{x})^{-1}-2p+2\right)}{\tau B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)} \end{cases}. \tag{8}$$

Theorem 1, which will be useful to introduce estimators of the tail index $\gamma(\mathbf{x})$ as part of our extrapolation methodology, generalizes and adapts to the conditional setup several results already found in the literature: see Theorem 1 in [5], Theorem 1 in [8] and Theorem 3 in [10]. Note however that, although they are in some sense related, Theorem 1 does not imply Theorem 1 of [17], because the latter is stated under weaker regularity conditions warranted by the specific context $p = 2$ of extreme conditional expectile estimation. On the technical side, assumptions (5) and (6) ensure that the bias introduced by smoothing in the $\mathbf{x}$ direction is negligible compared to the standard deviation σ_n of the estimator. The aim of the next paragraph is now to extrapolate our intermediate estimators to properly extreme levels.

3.2 Extreme L^p−quantile regression

We consider here a level $\beta_n \to 1$ such that $nh_n^d(1 - \beta_n) \to c < \infty$. The estimators previously introduced no longer work at such an extreme level. In order to overcome this problem, we first recall a result of [8] (see also Lemma 5 below):

$$\forall p \geq 1, \lim_{\alpha \to 1} \frac{q^{(p)}(\alpha|\mathbf{x})}{q^{(1)}(\alpha|\mathbf{x})} = \left(\frac{\gamma(\mathbf{x})}{B\left(p, \gamma(\mathbf{x})^{-1} - p + 1\right)}\right)^{-\gamma(\mathbf{x})}. \tag{9}$$

In the sequel we shall use the notation $g_p(\gamma) = \gamma/B\left(p, \gamma^{-1} - p + 1\right)$. A first consequence of this result is that the L^p−quantile function is regularly varying, i.e.:

$$\forall t > 0, \lim_{y \to \infty} \frac{q^{(p)}(1 - 1/(ty)|\mathbf{x})}{q^{(p)}(1 - 1/y|\mathbf{x})} = t^{\gamma(\mathbf{x})}. \tag{10}$$

This suggests then that, by considering an intermediate sequence (α_n), our conditional extreme L^p−quantile may be approximated (and estimated) as follows:

$$q^{(p)}(\beta_n|\mathbf{x}) \approx \left(\frac{1 - \alpha_n}{1 - \beta_n}\right)^{\gamma(\mathbf{x})} q^{(p)}(\alpha_n|\mathbf{x}),$$

$$\text{estimated by } \tilde{q}_{n,\alpha_n}^{(p)}(\beta_n|\mathbf{x}) = \left(\frac{1 - \alpha_n}{1 - \beta_n}\right)^{\hat{\gamma}_{\alpha_n}(\mathbf{x})} \hat{q}_n^{(p)}(\alpha_n|\mathbf{x}).$$

Here, $\hat{q}_n^{(p)}(\alpha_n|\mathbf{x})$ is the kernel estimator introduced in Equation (4) and $\hat{\gamma}_{\alpha_n}(\mathbf{x})$ is a consistent estimator of the conditional tail index $\gamma(\mathbf{x})$. This is a class of Weissman-type estimators (see [23]) of which we give the asymptotic properties.

Theorem 2 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and $\mathcal{C}_2(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(\cdot|\mathbf{x}))$ hold with $\rho(\mathbf{x}) < 0$. Let $\alpha_n, \beta_n \to 1$, $h_n \to 0$ be such that $\sigma_n^{-2} = nh_n^d(1 - \alpha_n) \to \infty$ and $nh_n^d(1 - \beta_n) \to c < \infty$. Assume further that $nh_n^{d+2}(1 - \alpha_n) \to 0$, $\omega_{h_n}^{(3)}(q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}) \to 0$ and

i) $\sigma_n^{-1}A\left((1 - \alpha_n)^{-1}|\mathbf{x}\right) = O(1)$, $\sigma_n^{-1}(1 - \alpha_n) = O(1)$ and
   $\sigma_n^{-1}\mathbb{E}\left[Y\mathbb{1}_{\{0 < Y < q^{(1)}(\alpha_n|\mathbf{x})\}}|\mathbf{x}\right]q^{(1)}(\alpha_n|\mathbf{x})^{-1} = O(1)$,
ii) For some $\delta \in (0, 1)$, $\sigma_n^{-1}\omega_{h_n}^{(1)}((1-\delta)[g_p(\gamma(\mathbf{x}))]^{-\gamma(\mathbf{x})}q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x})\log(1-\alpha_n) \to$
   0 and $\sigma_n^{-1}\omega_{h_n}^{(2)}((1 + \delta)q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}) \to 0$,
iii) $\sigma_n^{-1}/\log\left((1 - \alpha_n)/(1 - \beta_n)\right) \to \infty$.

Take $p \in (1, \gamma(\mathbf{x})^{-1}/2 + 1)$. If in addition $\sigma_n^{-1}(\hat{\gamma}_{\alpha_n}(\mathbf{x}) - \gamma(\mathbf{x})) \xrightarrow{d} \Gamma$, then

$$\frac{\sigma_n^{-1}}{\log((1 - \alpha_n)/(1 - \beta_n))} \left(\frac{\tilde{q}_{n,\alpha_n}^{(p)}(\beta_n|\mathbf{x})}{q^{(p)}(\beta_n|\mathbf{x})} - 1\right) \xrightarrow{d} \Gamma.$$

We notice, as is classical in the analysis of heavy tails, that the asymptotic distribution of the extrapolated estimator $\tilde{q}_{n,\alpha_n}^{(p)}(\beta_n|\mathbf{x})$ is exactly that of the purely empirical estimator $\hat{\gamma}_{\alpha_n}(\mathbf{x})$ with a slightly slower rate of convergence. Technically speaking, assumption (i) controls the bias due to the asymptotic approximation (9), while assumption (ii) is used to deal with the bias due to smoothing.

Our aim is now to propose some estimators of $\gamma(\mathbf{x})$ solely based on intermediate L^p−quantiles, in order to carry out the extrapolation step.

3.3 L^p−quantile based estimation of the conditional tail index

The aim of this paragraph is to discuss the estimation of the conditional tail index $\gamma(\mathbf{x})$. A local Pickands estimator is studied in [5, 6]. This estimator however has a large variance, which is why [6] propose a simplified, conditional and local version of the Hill estimator:

$$\hat{\gamma}_{\alpha_n}^{(H)}(\mathbf{x}) = \frac{1}{\log(J!)} \sum_{j=1}^{J} \log\left(\hat{q}_n\left(\frac{j-1+\alpha_n}{j}|\mathbf{x}\right) \Big/ \hat{q}_n\left(\alpha_n|\mathbf{x}\right)\right). \tag{11}$$

They also mentioned that taking $J = 9$ is an optimal choice, and leads to an asymptotic variance close to $1.25||K||_2^2 \gamma(\mathbf{x})^2/g(\mathbf{x})$. Recently, [9, 17] have shown that replacing the quantile by the expectile in tail index estimators can lead to a significant variance reduction. Our idea here is to propose an estimator based on L^p−quantiles rather than quantiles. In this context, we propose to follow the approach of [16] and exploit the asymptotic relationship (9) by introducing the following estimator, valid for all $1 < p < \gamma(\mathbf{x})^{-1} + 1$:

$$\hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x}) = \inf\left\{\gamma > 0 : g_p(\gamma) \le \frac{\hat{\bar{F}}_n^{(1)}\left(\hat{q}_n^{(p)}(\alpha_n|\mathbf{x})|\mathbf{x}\right)}{1 - \alpha_n}\right\}. \tag{12}$$

This class of estimators is introduced in [16] in an unconditional setting, and the (explicit) estimator $\hat{\gamma}_{\alpha_n}^{(2)}(\mathbf{x})$ is introduced in [17]. Using the results previously obtained, we can give the asymptotic distribution of $\hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x})$ for all $1 < p < \gamma(\mathbf{x})^{-1}/2 + 1$.

Theorem 3 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and $C_2(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(\cdot|\mathbf{x}))$ hold with $\gamma(\mathbf{x}) < 1$. Let $\alpha_n \to 1$ and $h_n \to 0$. Assume further that $\sigma_n^{-2} = nh_n^d(1 - \alpha_n) \to \infty$, $nh_n^{d+2}(1 - \alpha_n) \to 0$, $\omega_{h_n}^{(3)}(q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}) \to 0$ and

i) $\sigma_n^{-1} A\left((1 - \alpha_n)^{-1}|\mathbf{x}\right) \to 0$,
ii) $\sigma_n^{-1} q^{(1)}(\alpha_n|\mathbf{x})^{-1} \to \lambda \in \mathbb{R}$,
iii) For some $\delta \in (0, 1)$, $\sigma_n^{-1}\omega_{h_n}^{(1)}((1 - \delta)\left(g_p(\gamma(\mathbf{x}))^{-\gamma(\mathbf{x})} q^{(1)}(\alpha_n|\mathbf{x})\right)|\mathbf{x})\log(1 - \alpha_n) \to 0$ and $\sigma_n^{-1}\omega_{h_n}^{(2)}((1 + \delta)\left(q^{(1)}(\alpha_n|\mathbf{x})\right)|\mathbf{x}) \to 0$.

Then, for all $p \in (1, \gamma(\mathbf{x})^{-1}/2 + 1)$, one has

$$\sigma_n^{-1}\left(\hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x}) - \gamma(\mathbf{x}), \frac{\hat{q}_n^{(p)}(\alpha_n|\mathbf{x})}{q^{(p)}(\alpha_n|\mathbf{x})} - 1\right) \overset{d}{\longrightarrow} \boldsymbol{\Theta}, \tag{13}$$

where $\boldsymbol{\Theta}$ is a bivariate Gaussian distribution with mean vector $\left(b_p(\mathbf{x}), 0\right)$ and covariance matrix $||K||_2^2 \gamma(\mathbf{x})^2 g(\mathbf{x})^{-1}\boldsymbol{\Omega}(\mathbf{x})$ such that:

$$
\begin{cases}
b_p(\mathbf{x}) = \frac{(1-p)\gamma(\mathbf{x})g_p(\gamma(\mathbf{x}))^{\gamma(\mathbf{x})}\mathbb{E}[Y|\mathbf{X=x}]}{1-\frac{1}{\gamma(\mathbf{x})}\left(\Psi(\gamma(\mathbf{x})^{-1}+1)-\Psi(\gamma(\mathbf{x})^{-1}-p+1)\right)}\lambda \\[2ex]
\Omega_{11}(\mathbf{x}) = \frac{B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)}{\left(1-\frac{1}{\gamma(\mathbf{x})}\left(\Psi(\gamma(\mathbf{x})^{-1}+1)-\Psi(\gamma(\mathbf{x})^{-1}-p+1)\right)\right)^2}\left(\frac{B\left(2p-1,\gamma(\mathbf{x})^{-1}-2p+2\right)}{B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)^2}-\frac{1}{\gamma(\mathbf{x})}\right) \\[2ex]
\Omega_{12}(\mathbf{x}) = \frac{B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)}{1-\frac{1}{\gamma(\mathbf{x})}\left(\Psi(\gamma(\mathbf{x})^{-1}+1)-\Psi(\gamma(\mathbf{x})^{-1}-p+1)\right)}\left(\frac{1}{\gamma(\mathbf{x})}-\frac{B\left(2p-1,\gamma(\mathbf{x})^{-1}-2p+2\right)}{B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)^2}\right) \\[2ex]
\Omega_{22}(\mathbf{x}) = \frac{B\left(2p-1,\gamma(\mathbf{x})^{-1}-2p+2\right)}{B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)}
\end{cases}.
$$

$$(14)$$

Let us remark here that although Theorem 3 can be seen as a version of Theorem 4 of [17], the latter is stated under weaker regularity assumptions and applies to further examples of estimators developed specifically in the conditional expectile setup.

Note that condition $\gamma(\mathbf{x}) < 1$ entails $\mathbb{E}[Y|\mathbf{X} = \mathbf{x}] < \infty$ and leads to a simple expression of the bias term $b_p(\mathbf{x})$. A result dropping this assumption is available in the unconditional setting in [16]; here, our motivation for this condition is that we shall use extreme regression L^p−quantiles as a way to estimate extreme regression expectiles, for the existence of which a natural condition is that $\mathbb{E}[|Y||\mathbf{X} = \mathbf{x}] < \infty$.

The bias term $b_p(\mathbf{x})$ is related to $\gamma(\mathbf{x})$, $q^{(1)}(\alpha_n|\mathbf{x})$ and $\mathbb{E}[Y|\mathbf{X} = \mathbf{x}]$. All these quantities may be easily estimated (the latter two by kernel regression estimators) to construct a bias-reduced conditional tail index estimator as follows:

$$
\tilde{\gamma}_{\alpha_n}^{(p)}(\mathbf{x}) = \hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x})\left(1+\frac{(p-1)\left(\frac{\sum_{i=1}^n Y_i K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\right)}\right)\hat{q}_n^{(p)}(\alpha_n|\mathbf{x})^{-1}}{1+\frac{1}{\hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x})}\left(\Psi\left(1/\hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x})-p+1\right)-\Psi\left(1/\hat{\gamma}_{\alpha_n}^{(p)}(\mathbf{x})+1\right)\right)}\right).
$$

Under the conditions of Theorem 3, it is clear that $\sigma_n^{-1}(\tilde{\gamma}_{\alpha_n}^{(p)}(\mathbf{x}) - \gamma(\mathbf{x})) \xrightarrow{d} \mathcal{N}(0,\Omega_{11}(\mathbf{x}))$ where $\Omega_{11}(\mathbf{x})$ is given in Equation (14). This bias reduction improves significantly the numerical results, and is used in the finite-sample study below.

Even though L^p−quantiles with $1 < p < 2$ are more widely estimable than expectiles and take into account the whole tail information, they are neither easy to interpret nor coherent as risk measures. Recent work in [8] has shown that extreme L^p−quantiles can be used as vehicles for extreme quantile and expectile estimation; see also [15] for an analogous study of the estimation of (a compromise between) Median Shortfall and Conditional Tail Expectation at extreme levels, using tail L^p−medians. Our focus in the following finite-sample study is to analyse the potential of extreme regression L^p−quantiles for the estimation of extreme regression quantiles and expectiles.

4 Simulation study

We consider here a one-dimensional covariate ($d = 1$), uniformly distributed on
$[0, 1]$, and a Burr-type distribution for Y given $X = x$:

$$\bar{F}^{(1)}(y|x) = \left(1 + y^{-\rho(x)/\gamma(x)}\right)^{1/\rho(x)}, \ \gamma(x) = \frac{4 + \sin(2\pi x)}{10} \text{ and } \rho(x) \equiv -1.$$

Such a distribution fulfills Assumption $C_2(\gamma(x), \rho(x), A(\cdot|x))$ with auxiliary func-
tion $A(y|x) = \gamma(x)y^{\rho(x)}$. We simulate $N = 500$ samples of size $n = 1,000$ indepen-
dent replications of (X, Y), and propose to estimate the conditional quantiles and
expectiles of level $\beta_n = 1 - 1/n = 0.999$ using our extreme regression L^p-quantile
estimators. Note that the quantiles may be calculated explicitly:

$$q(\alpha|x) = \left[(1 - \alpha)^{\rho(x)} - 1\right]^{-\gamma(x)/\rho(x)}.$$

Expectiles have to be approximated numerically, since they do not have a simple
closed form. In order to estimate these two quantities, we propose to compare
different approaches (called either direct or indirect):

(i) Use the conditional Weissman-type estimators respectively based on empirical
quantiles and the estimator $\hat{\gamma}_{\alpha_n}^{(H)}(x)$ (direct quantile estimator) and on empirical
expectiles and $\tilde{\gamma}_{\alpha_n}^{(2)}(x)$ (direct expectile estimator), i.e.

$$\left(\frac{1 - \alpha_n}{1 - \beta_n}\right)^{\hat{\gamma}_{\alpha_n}^{(H)}(x)} \hat{q}_n^{(1)}(\alpha_n|x) \, , \ \left(\frac{1 - \alpha_n}{1 - \beta_n}\right)^{\tilde{\gamma}_{\alpha_n}^{(2)}(x)} \hat{q}_n^{(2)}(\alpha_n|x).$$

(ii) Indirect quantile estimator: estimate first the conditional L^p-quantile using
estimator (4), and exploit asymptotic relationship (9) to recover the extreme
conditional quantile,

$$\left(\frac{1 - \alpha_n}{1 - \beta_n}\right)^{\tilde{\gamma}_{\alpha_n}^{(p)}(x)} \hat{q}_n^{(p)}(\alpha_n|x) \left(\frac{\tilde{\gamma}_{\alpha_n}^{(p)}(x)}{B\left(p, \tilde{\gamma}_{\alpha_n}^{(p)}(x)^{-1} - p + 1\right)}\right)^{\tilde{\gamma}_{\alpha_n}^{(p)}(x)}.$$

(iii) Indirect expectile estimator: use Equation (9) to get a connection between
L^p-quantile and quantile, and quantile and expectile, resulting in the extreme
conditional expectile estimator

$$\left(\frac{1 - \alpha_n}{1 - \beta_n}\right)^{\tilde{\gamma}_{\alpha_n}^{(p)}(x)} \hat{q}_n^{(p)}(\alpha_n|x) \left(\frac{B\left(2, \tilde{\gamma}_{\alpha_n}^{(p)}(x)^{-1} - 1\right)}{B\left(p, \tilde{\gamma}_{\alpha_n}^{(p)}(x)^{-1} - p + 1\right)}\right)^{\tilde{\gamma}_{\alpha_n}^{(p)}(x)}.$$

The choice of p is discussed in [16] using the MSE of (the unconditional version of)
$\tilde{\gamma}_{\alpha_n}^{(p)}(x)$ as a criterion. Cross-validation choices of the bandwidth h_n and intermediate
quantile level α_n, meanwhile, are discussed [5, 17]. For the sake of simplicity, we
choose here common parameters $p = 1.7$ following the guidelines of [16]), $h_n = 0.15$

and $\alpha_n = 1 - 1/\sqrt{n} \approx 0.968$ across all replications and K is the Epanechnikov kernel defined by $K(t) = 0.75(1 - t^2)\mathbb{1}_{\{|t|<1\}}$. Results are shown in Figure 1.
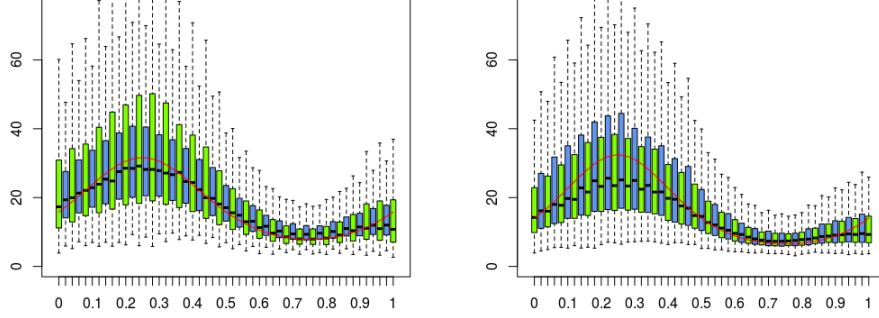

Fig. 1 Left: Boxplots of 500 estimates of $q^{(1)}(\beta_n|x)$ with the direct (green) and indirect (blue) quantile estimators. Right: Boxplots of 500 estimates of $q^{(2)}(\beta_n|x)$ with the direct (green) and indirect (blue) expectile estimators. True values are in red.

We can notice that an indirect estimation of extreme quantiles or expectiles with a L^p−quantile (with p between 1 and 2) leads to a trade-off between bias and variance: the indirect L^p−estimator of an extreme regression quantile is less variable than the direct estimator but slightly more biased, and the indirect L^p−estimator of an extreme regression expectile is more variable than the direct estimator but less biased. For conditional quantiles, an explanation is that using the asymptotic approximation (9) in the construction of the indirect estimator adds a source of bias, while the reduced variance stems from the use of $p = 1.7$ in the estimator $\tilde{\gamma}_{\alpha_n}^{(p)}(x)$, providing an estimator with lower variance compared to the simple Hill estimator in our case (see [16]). The case of conditional expectiles is less clear, although the increased variability observed for $x \in [0, 0.5]$ seems to originate in the use of the estimated constant $B(2, \tilde{\gamma}_{\alpha_n}^{(p)}(x)^{-1} - 1)/B(p, \tilde{\gamma}_{\alpha_n}^{(p)}(x)^{-1} - p + 1)$: when $\tilde{\gamma}_{\alpha_n}^{(p)}(x)$ gets close to 1, which is sometimes the case in this zone where $\gamma(x) \in [0.4, 0.5]$, this estimated constant tends to explode, while the direct estimator is less affected. A similar observation, in the context of extreme Wang distortion risk measure estimation, is made by [13].

5 Real data example

We study here a data set on motorcycle insurance, collected from the former Swedish insurance provider Wasa. Our data is on motorcycle insurance policies and claims over the period 1994-1998 and is available from www.math.su.se/GLMbook or the R packages `insuranceData` and `CASdatasets`, and analyzed in [21]. We concentrate here on the relationship between the claim severity Y (defined as the

ratio of claim cost by number of claims for each given policyholder) in Swedish kroner (SEK), and the number of years X of exposure of a policyholder. Data for $X > 3$ are very sparse, so we restrict our attention to the case $Y > 0$ and $X \in [0, 3]$, resulting in $n = 593$ pairs (X_i, Y_i).

Our goal in this section is to estimate extreme conditional quantiles and expectiles of Y given X, at a level $\beta_n = 1 - 3/n \approx 0.9949$. This level is slightly less extreme than the more standard $\beta_n = 1 - 1/n \approx 0.9985$, but is an appropriately extreme level in this conditional context where less data are available locally for the estimation. A preliminary diagnostic using a local version of the Hill estimator (which we do not show here) suggests that the data is indeed heavy-tailed with $\gamma(x) \in [0.25, 0.6]$. Following again the guidelines in [16], we choose $p = 1.7$ for our indirect extreme conditional quantile and expectile estimators. These are respectively compared to:

- the estimator $\widehat{q}_n^W(\beta_n|x)$ of [17], calculated as in Section 5 therein, and our direct quantile estimator presented in Section 4 (i),
- the estimator $\widehat{e}_n^{W,BR}(\beta_n|x)$ of [17], calculated as in Section 5 therein, and our direct expectile estimator presented in Section 4 (i).

For the direct and indirect estimators presented in Section 4 (ii)-(iii), the parameters α_n and h_n are chosen by a cross-validation procedure analogous to that of [17]. The Epanechnikov kernel is adopted. Results are given in Figure 2. In each case, all three estimators reassuringly point to roughly the same results, with slight differences; in particular, for quantile estimation and when data is scarce, the direct estimator in Section 4 (i) appears to be more sensitive to the local shape of the tail than the indirect, L^p−quantile based estimator in Section 4 (ii), resulting in less stable estimates.

6 Appendix

6.1 Preliminary results

Lemma 1 Assume that $(\mathcal{L})$ and $\mathcal{C}_2$ $(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ hold, and let $y_n \to \infty$ and $h_n \to 0$ be such that $\omega_{h_n}^{(1)}(y_n|\mathbf{x}) \log(y_n) \to 0$ and $\omega_{h_n}^{(2)}(y_n|\mathbf{x}) \to 0$. Then for all $0 \le k < \gamma(\mathbf{x})^{-1}$ we have, uniformly in $\mathbf{x}' \in B(\mathbf{x}, h_n)$,

$$m^{(k)}(y_n|\mathbf{x}') = m^{(k)}(y_n|\mathbf{x}) \left(1 + O(h_n) + o\left(\omega_{h_n}^{(1)}(y_n|\mathbf{x})\right) + O\left(\omega_{h_n}^{(2)}(y_n|\mathbf{x})\right)\right).$$

In particular $m^{(k)}(y_n|\mathbf{x}') = y_n^k g(\mathbf{x}) (1 + o(1))$ uniformly in $\mathbf{x}' \in B(\mathbf{x}, h_n)$.

Proof Let us first write

$$m^{(k)}(y_n|\mathbf{x}) = \mathbb{E}\left[(Y - y_n)^k \mathbb{1}_{\{Y > y_n\}} | \mathbf{X} = \mathbf{x}\right] g(\mathbf{x}) + \mathbb{E}\left[(y_n - Y)^k \mathbb{1}_{\{Y \le y_n\}} | \mathbf{X} = \mathbf{x}\right] g(\mathbf{x}).$$

By the arguments of the proof of Lemma 3 in [17],

$$\frac{\mathbb{E}\left[(Y - y_n)^k \mathbb{1}_{\{Y > y_n\}} | \mathbf{X} = \mathbf{x}'\right] g(\mathbf{x}')}{\mathbb{E}\left[(Y - y_n)^k \mathbb{1}_{\{Y > y_n\}} | \mathbf{X} = \mathbf{x}\right] g(\mathbf{x})} = 1 + O(h_n) + O\left(\omega_{h_n}^{(1)}(y_n|\mathbf{x}) \log(y_n)\right).$$
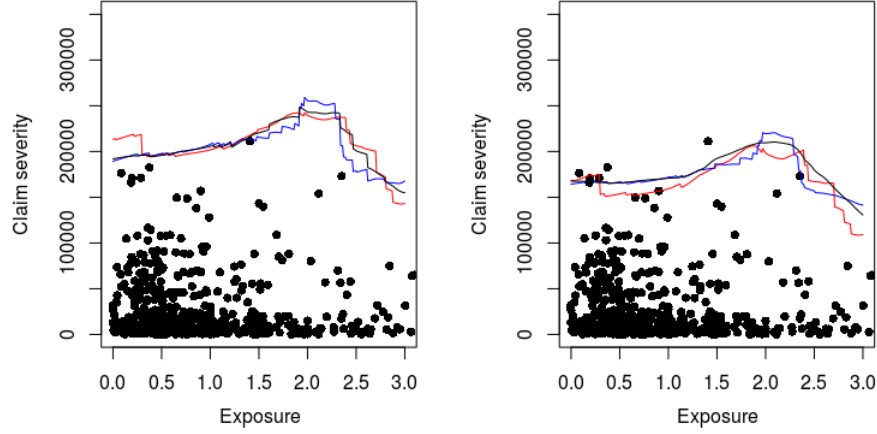
Fig. 2 Swedish motorcycle insurance data. Left panel: extreme conditional quantile estimation, black curve: estimator $\widehat{q}_n^W(\beta_n|x)$ of [17], blue curve: direct quantile estimator (i) of Section 4, red curve: indirect quantile estimator (ii) of Section 4. Right panel: extreme conditional expectile estimation, black curve: estimator $\widehat{e}_n^{W,BR}(\beta_n|x)$ of [17], blue curve: direct expectile estimator (i) of Section 4, red curve: indirect expectile estimator (iii) of Section 4. In each panel, x-axis: number of years of exposure of policyholder, y-axis: claim severity.

Besides, an integration by parts yields

$$\mathbb{E}\left[(y_n - Y)^k \mathbb{1}_{\{Y \le y_n\}}|\mathbf{X} = \mathbf{x}\right] = \int_0^{y_n} k t^{k-1} F^{(1)}(y_n - t|\mathbf{x})\, dt.$$

It clearly follows that

$$\left|\mathbb{E}\left[(y_n - Y)^k \mathbb{1}_{\{Y \le y_n\}}|\mathbf{X} = \mathbf{x}'\right] - \mathbb{E}\left[(y_n - Y)^k \mathbb{1}_{\{Y \le y_n\}}|\mathbf{X} = \mathbf{x}\right]\right| \le y_n^k \omega_{h_n}^{(2)}(y_n|\mathbf{x}).$$

Now

$$\mathbb{E}\left[(y_n - Y)^k \mathbb{1}_{\{Y \le y_n\}}|\mathbf{X} = \mathbf{x}\right] = y_n^k \mathbb{E}\left[\left(1 - \frac{Y}{y_n}\right)^k \mathbb{1}_{\{Y \le y_n\}}|\mathbf{X} = \mathbf{x}\right] = y_n^k (1 + o(1))$$

by the dominated convergence theorem, and

$$\mathbb{E}\left[(Y - y_n)^k \mathbb{1}_{\{Y > y_n\}}|\mathbf{X} = \mathbf{x}\right] = \frac{g(\mathbf{x}) B\left(k + 1, \gamma(\mathbf{x})^{-1} - k\right)}{\gamma(\mathbf{x})} y_n^k \bar{F}^{(1)}(y_n|\mathbf{x})(1 + o(1)),$$

$$(15)$$

see for instance Lemma 1(i) in [8]. The result follows from direct calculations.

Lemma 2 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and $\mathcal{C}_2(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ hold, and let $y_n \to \infty$ and $h_n \to 0$ be such that $nh_n^d \to \infty$, $\omega_{h_n}^{(1)}(y_n|\mathbf{x}) \log(y_n) \to 0$ and $\omega_{h_n}^{(2)}(y_n|\mathbf{x}) \to 0$. Then for all $0 \le k < \gamma(\mathbf{x})^{-1}/2$,

$$\mathbb{E}\left[\hat{m}_n^{(k)}(y_n|\mathbf{x})\right] = m^{(k)}(y_n|\mathbf{x})\left(1 + O(h_n) + o\left(\omega_{h_n}^{(1)}(y_n|\mathbf{x})\right) + O\left(\omega_{h_n}^{(2)}(y_n|\mathbf{x})\right)\right)$$

and $\mathbb{V}\mathrm{ar}\left[\hat{m}_n^{(k)}(y_n|\mathbf{x})\right] = \dfrac{||K||_2^2}{nh_n^d}g(\mathbf{x})y_n^{2k}(1 + o(1))$. $\hfill\square$

Proof Note that $\mathbb{E}\left[\hat{m}_n^{(k)}(y_n|\mathbf{x})\right] = \int_S m^{(k)}(y_n|\mathbf{x}-\mathbf{u}h_n)K(\mathbf{u})d\mathbf{u}$ by Assumption $(\mathcal{K})$ and a change of variables, and use Lemma 1 to get the first result. The second result is obtained through similar calculations. $\hfill\square$

Lemma 3 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and $C_2\left(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x})\right)$ hold. Let $y_n \to \infty$, $h_n \to 0$ be such that $nh_n^d \to \infty$ and $\omega_{h_n}^{(1)}(y_n|\mathbf{x})\log(y_n) \to 0$. Then for all $0 \le k < \gamma(\mathbf{x})^{-1}/2$,

$$\begin{cases} \mathbb{E}\left[\hat{\varphi}_n^{(k)}(y_n|\mathbf{x})\right] & = \varphi^{(k)}(y_n|\mathbf{x})\left(1 + O(h_n) + O\left(\omega_{h_n}^{(1)}(y_n|\mathbf{x})\log(y_n)\right)\right), \\ \mathbb{V}\mathrm{ar}\left[\hat{\varphi}_n^{(k)}(y_n|\mathbf{x})\right] & = ||K||_2^2 g(\mathbf{x})\dfrac{B\left(2k+1, \gamma(\mathbf{x})^{-1}-2k\right)}{\gamma(\mathbf{x})}\dfrac{y_n^{2k}\bar{F}^{(1)}(y_n|\mathbf{x})}{nh_n^d}(1 + o(1)). \end{cases}$$

Proof See Lemma 5 of [17]. $\hfill\square$

Lemma 4 Assume that $C_2\left(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x})\right)$ holds. Let $\lambda \ge 1$, $y_n \to \infty$, $y_n' = \lambda y_n(1 + o(1))$ and $0 < k < \gamma(\mathbf{x})^{-1}$.

(i) Then the following asymptotic relationship holds:

$$\mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y>y_n'\}}|\mathbf{X} = \mathbf{x}\right]$$
$$= y_n^k \bar{F}^{(1)}(y_n|\mathbf{x})\left[kIB\left(\lambda^{-1}, \gamma(\mathbf{x})^{-1} - k, k\right) + (\lambda - 1)^k\lambda^{-1/\gamma(\mathbf{x})}\right](1 + o(1)). \quad\square$$

(ii) Assume further that $\omega_{h_n}^{(1)}(y_n \wedge y_n'|\mathbf{x})\log(y_n) \to 0$ and $\omega_{h_n}^{(3)}(y_n|\mathbf{x}) \to 0$. Then, uniformly in $\mathbf{x}' \in B(\mathbf{x}, h_n)$,

$$\mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y>y_n'\}}|\mathbf{X} = \mathbf{x}'\right] = \mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y>y_n'\}}|\mathbf{X} = \mathbf{x}\right](1 + o(1)).$$

Proof (i) Straightforward calculations entail

$$\mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y>y_n'\}}|\mathbf{X} = \mathbf{x}\right]$$
$$= y_n^k \mathbb{E}\left[\left\{\left(\frac{Y}{y_n} - 1\right)^k - (\lambda - 1)^k\right\}\mathbb{1}_{\{Y>\lambda y_n\}}|\mathbf{X} = \mathbf{x}\right](1 + o(1))$$
$$+ y_n^k(\lambda - 1)^k \bar{F}^{(1)}(\lambda y_n|\mathbf{x})(1 + o(1)),$$

with $y_n' = \lambda y_n(1 + o(1))$. The result then comes directly from the regular variation property of $\bar{F}^{(1)}(\cdot|\mathbf{x})$ and Lemma 1 in [8] with $H(t) = (t - 1)^k$ and $b = \lambda$.

(ii) Note first that for n large enough

$$\left| \mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y > y'_n\}}|\mathbf{X} = \mathbf{x}'\right] - \mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y > \lambda y_n\}}|\mathbf{X} = \mathbf{x}'\right] \right|$$

$$\leq \left[|y'_n - y_n|^k + (\lambda - 1)^k y_n^k\right]\left[\bar{F}^{(1)}(y'_n \wedge \lambda y_n|\mathbf{x}') - \bar{F}^{(1)}(y'_n \vee \lambda y_n|\mathbf{x}')\right]$$

$$\leq 3(\lambda - 1)^k y_n^k \times \bar{F}^{(1)}(y'_n|\mathbf{x}') \times \omega_{h_n}^{(3)}(y_n|\mathbf{x}).$$

Write $(Y - y_n)^k = ((Y - y_n)^k - (\lambda - 1)^k y_n^k) + (\lambda - 1)^k y_n^k$. It then follows from the assumption $\omega_{h_n}^{(3)}(y_n|\mathbf{x}) \to 0$ that, uniformly in $\mathbf{x}' \in B(\mathbf{x}, h_n)$,

$$\mathbb{E}\left[|Y - y_n|^k \mathbb{1}_{\{Y > y'_n\}}|\mathbf{X} = \mathbf{x}'\right] = (\lambda - 1)^k y_n^k \bar{F}^{(1)}(y'_n|\mathbf{x}')(1 + o(1))$$

$$+ k \int_{\lambda y_n}^{\infty} (z - y_n)^{k-1} \bar{F}^{(1)}(z|\mathbf{x}')dz(1 + o(1)).$$

Remark now $\bar{F}^{(1)}(y'_n|\mathbf{x})(y'_n)^{-\omega_{h_n}^{(1)}(y'_n|\mathbf{x})} \leq \bar{F}^{(1)}(y'_n|\mathbf{x}') \leq \bar{F}^{(1)}(y'_n|\mathbf{x})(y'_n)^{\omega_{h_n}^{(1)}(y'_n|\mathbf{x})}$. Then condition $\omega_{h_n}^{(1)}(y'_n|\mathbf{x})\log(y_n) \to 0$ entails, uniformly in $\mathbf{x}' \in B(\mathbf{x}, h_n)$, $\bar{F}^{(1)}(y'_n|\mathbf{x}') = \bar{F}^{(1)}(y'_n|\mathbf{x})(1 + o(1)) = \bar{F}^{(1)}(\lambda y_n|\mathbf{x})(1 + o(1))$. Besides, for any $z \geq \lambda y_n \geq y_n$, $\bar{F}^{(1)}(z|\mathbf{x})z^{-\omega_{h_n}^{(1)}(y_n|\mathbf{x})} \leq \bar{F}^{(1)}(z|\mathbf{x}') \leq \bar{F}^{(1)}(z|\mathbf{x})z^{\omega_{h_n}^{(1)}(y_n|\mathbf{x})}$. Following the proof of Lemma 3 in [17], we get, uniformly in $\mathbf{x}' \in B(\mathbf{x}, h_n)$,

$$\left| \frac{\int_{\lambda y_n}^{\infty} (z - y_n)^{k-1} \bar{F}^{(1)}(z|\mathbf{x}')dz}{\int_{\lambda y_n}^{\infty} (z - y_n)^{k-1} \bar{F}^{(1)}(z|\mathbf{x})dz} - 1 \right| = O(\omega_{h_n}^{(1)}(y_n|\mathbf{x})\log(y_n)) \to 0.$$

Since $\int_{\lambda y_n}^{\infty} (z - y_n)^{k-1} \bar{F}^{(1)}(z|\mathbf{x})dz$ is of order $y_n^k \bar{F}^{(1)}(y_n|\mathbf{x})$ (by regular variation of $\bar{F}^{(1)}(\cdot|\mathbf{x})$), the conclusion follows. $\qquad \square$

Lemma 5 Assume that $C_2(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ holds. For all $1 \leq p < \gamma(\mathbf{x})^{-1} + 1$,

$$\frac{\bar{F}^{(p)}(y|\mathbf{x})}{\bar{F}^{(1)}(y|\mathbf{x})} = \frac{B\left(p, \gamma(\mathbf{x})^{-1} - p + 1\right)}{\gamma(\mathbf{x})}\left[1 + r(y|\mathbf{x})\right]$$

where there are constants $C_1(\mathbf{x}), C_2(\mathbf{x}), C_3(\mathbf{x})$ such that

$$r(y|\mathbf{x}) = C_1(\mathbf{x})\frac{\mathbb{E}(Y\mathbb{1}_{\{0 < Y < y\}}|\mathbf{X} = \mathbf{x})}{y}(1 + o(1)) + C_2(\mathbf{x})\bar{F}^{(1)}(y|\mathbf{x})(1 + o(1))$$

$$+ C_3(\mathbf{x})A(1/\bar{F}^{(1)}(y|\mathbf{x})|\mathbf{x})(1 + o(1)) \text{ as } y \to \infty.$$

Similarly

$$\frac{q^{(p)}(\alpha|\mathbf{x})}{q^{(1)}(\alpha|\mathbf{x})} = \left(\frac{\gamma(\mathbf{x})}{B\left(p, \gamma(\mathbf{x})^{-1} - p + 1\right)}\right)^{-\gamma(\mathbf{x})}\left[1 + R(\alpha|\mathbf{x})\right]$$

where there are constants $D_1(\mathbf{x}), D_2(\mathbf{x}), D_3(\mathbf{x})$ such that

$$R(\alpha|\mathbf{x}) = D_1(\mathbf{x})\frac{\mathbb{E}(Y\mathbb{1}_{\{0<Y<q^{(1)}(\alpha|\mathbf{x})\}}|\mathbf{X}=\mathbf{x})}{q^{(1)}(\alpha|\mathbf{x})}(1+o(1)) + D_2(\mathbf{x})(1-\alpha)(1+o(1))$$

$$+ D_3(\mathbf{x})A((1-\alpha)^{-1}|\mathbf{x})(1+o(1)) \quad \text{as } \alpha \to 1. \qquad \qquad \square$$

Proof We start by focusing on the ratio $\bar{F}^{(p)}(y|\mathbf{x})/\bar{F}^{(1)}(y|\mathbf{x})$. By Lemma 1 in [16], the function $\bar{F}^{(p)}(\cdot|\mathbf{x})$ is continuous and strictly decreasing on the support of Y given $\mathbf{X} = \mathbf{x}$. It is therefore enough to show the announced formula for $y = q^{(p)}(\alpha|\mathbf{x})$ with $\alpha \to 1$; this, in turn, is a simple corollary of Proposition 2 in [8]. To show the analogous formula on $q^{(p)}(\alpha|\mathbf{x})/q^{(1)}(\alpha|\mathbf{x})$, we define $U^{(1)}(t|\mathbf{x}) = q^{(1)}(1 - t^{-1}|\mathbf{x})$; $U^{(1)}(\cdot|\mathbf{x})$ also satisfies a (local uniform) second-order regular variation condition, see Theorem 2.3.9 p.48 in [11]. Consequently, we note that the asymptotic expansion on $\bar{F}^{(p)}(y|\mathbf{x})/\bar{F}^{(1)}(y|\mathbf{x})$ entails a similar expansion on

$$\frac{U^{(1)}(1/\bar{F}^{(1)}(y|\mathbf{x})|\mathbf{x})}{U^{(1)}(1/\bar{F}^{(p)}(y|\mathbf{x})|\mathbf{x})} = \frac{y}{q^{(1)}(F^{(p)}(y|\mathbf{x}))}(1+o(A(1/\bar{F}^{(1)}(y|\mathbf{x})|\mathbf{x})))$$

as $y \to \infty$, with different constants (here Lemma 1 in [10] was used). Setting $y = q^{(p)}(\alpha|\mathbf{x})$, with $\alpha \to 1$, gives the announced result. $\qquad \qquad \square$

Lemma 6 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and $C_2(\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ hold. Let $y_n \to \infty$, $h_n \to 0$ and $z_n = \theta y_n(1 + o(1))$, where $\theta > 0$. Assume further that $\epsilon_n^{-2} = nh_n^d\bar{F}^{(1)}(y_n|\mathbf{x}) \to \infty$, $nh_n^{d+2}\bar{F}^{(1)}(y_n|\mathbf{x}) \to 0$, there exists $\delta \in (0,1)$ such that $\epsilon_n^{-1}\omega_{h_n}^{(1)}((1-\delta)(\theta \wedge 1)y_n|\mathbf{x})\log(y_n) \to 0$, and $\omega_{h_n}^{(3)}(z_n|\mathbf{x}) \to 0$. Letting, for all $j \in \{1, \ldots, J\}$, $y_{n,j} = \tau_j^{-\gamma(\mathbf{x})}y_n(1 + o(1))$ with $0 < \tau_1 < \tau_2 < \ldots < \tau_J \le 1$, and $p \in (1, \gamma(\mathbf{x})^{-1}/2 + 1)$, one has

$$\epsilon_n^{-1}\left\{\left(\frac{\hat{\varphi}_n^{(0)}(y_{n,j}|\mathbf{x})}{\varphi^{(0)}(y_{n,j}|\mathbf{x})} - 1\right)_{1 \le j \le J}, \left(\frac{\hat{\varphi}_n^{(p-1)}(z_n|\mathbf{x})}{\varphi^{(p-1)}(z_n|\mathbf{x})} - 1\right)\right\} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}_{J+1}, \frac{\|K\|_2^2}{g(\mathbf{x})}\mathbf{\Lambda}(\mathbf{x})\right),$$

where $\mathbf{\Lambda}(\mathbf{x})$ is a symmetric matrix having entries:

$$\begin{cases} \Lambda_{j,\ell}(\mathbf{x}) = & \left(\tau_j \vee \tau_\ell\right)^{-1} \\ \Lambda_{j,J+1}(\mathbf{x}) = & \gamma(\mathbf{x})\dfrac{(p-1)IB\left(\left(1\vee\frac{\tau_j^{-\gamma(\mathbf{x})}}{\theta}\right)^{-1}, \gamma(\mathbf{x})^{-1}-p+1, p-1\right) + \left(1\vee\frac{\tau_j^{-\gamma(\mathbf{x})}}{\theta}-1\right)^{p-1}\left(1\vee\frac{\tau_j^{-\gamma(\mathbf{x})}}{\theta}\right)^{-1/\gamma(\mathbf{x})}}{\tau_j B\left(p, \gamma(\mathbf{x})^{-1}-p+1\right)} \\ \Lambda_{J+1,J+1}(\mathbf{x}) = & \gamma(\mathbf{x})\dfrac{B\left(2p-1, \gamma(\mathbf{x})^{-1}-2p+2\right)}{B\left(p, \gamma(\mathbf{x})^{-1}-p+1\right)^2}\theta^{1/\gamma(\mathbf{x})} \end{cases}.$$

$$\tag{16}$$

Proof Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J, \beta_{J+1}) \in \mathbb{R}^{J+1}$. Set

$$\mathcal{Z}_n = \epsilon_n^{-1}\sum_{j=1}^J \beta_j\left(\frac{\hat{\varphi}_n^{(0)}(y_{n,j}|\mathbf{x})}{\varphi^{(0)}(y_{n,j}|\mathbf{x})} - 1\right) + \epsilon_n^{-1}\beta_{J+1}\left(\frac{\hat{\varphi}_n^{(p-1)}(z_n|\mathbf{x})}{\varphi^{(p-1)}(z_n|\mathbf{x})} - 1\right).$$

Clearly $\omega_{h_n}^{(1)}(y_{n,j}|\mathbf{x}) \le \omega_{h_n}^{(1)}((1-\delta)y_n|\mathbf{x})$ and $\omega_{h_n}^{(1)}(z_n|\mathbf{x}) \le \omega_{h_n}^{(1)}((1-\delta)\theta y_n|\mathbf{x})$ for n large enough. Lemma 3 then provides $\mathbb{E}(\mathcal{Z}_n) = o(1)$. It thus remains to focus on the asymptotic distribution of the centered variable $Z_n = \mathcal{Z}_n - \mathbb{E}(\mathcal{Z}_n)$. Note that $\mathbb{V}\text{ar}[Z_n] = \epsilon_n^{-2}\boldsymbol{\beta}^\top \boldsymbol{B}^{(n)}\boldsymbol{\beta}$, where $\boldsymbol{B}^{(n)}$ is the symmetric matrix having entries:

$$
\begin{cases}
B_{j,\ell}^{(n)}(\mathbf{x}) = \dfrac{\text{cov}\left(\hat{\varphi}_n^{(0)}(y_{n,j}|\mathbf{x}),\hat{\varphi}_n^{(0)}(y_{n,\ell}|\mathbf{x})\right)}{\varphi^{(0)}(y_{n,j}|\mathbf{x})\varphi^{(0)}(y_{n,\ell}|\mathbf{x})}, & j,\ell \in \{1,\ldots,J\},\ j \le \ell, \\[2ex]
B_{j,J+1}^{(n)}(\mathbf{x}) = \dfrac{\text{cov}\left(\hat{\varphi}_n^{(0)}(y_{n,j}|\mathbf{x}),\hat{\varphi}_n^{(p-1)}(z_n|\mathbf{x})\right)}{\varphi^{(0)}(y_{n,j}|\mathbf{x})\varphi^{(p-1)}(z_n|\mathbf{x})}, & j \in \{1,\ldots,J\}, \\[2ex]
B_{J+1,J+1}^{(n)}(\mathbf{x}) = \dfrac{\mathbb{V}\text{ar}\left[\hat{\varphi}_n^{(p-1)}(z_n|\mathbf{x})\right]}{\varphi^{(p-1)}(z_n|\mathbf{x})^2}.
\end{cases}
$$

We recall $z_n = \theta y_n(1+o(1))$, hence $\bar{F}^{(1)}(z_n|\mathbf{x}) = \theta^{-1/\gamma(\mathbf{x})}\bar{F}^{(1)}(y_n|\mathbf{x})(1+o(1))$ and Lemma 3 combined with Equation (15) immediately gives

$$
B_{J+1,J+1}^{(n)}(\mathbf{x}) = \frac{||K||_2^2}{g(\mathbf{x})}\gamma(\mathbf{x})\frac{B\left(2p-1,\gamma(\mathbf{x})^{-1}-2p+2\right)}{B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)^2}\theta^{1/\gamma(\mathbf{x})}\epsilon_n^2(1+o(1)).
$$

The calculation of $B_{j,\ell}^{(n)}(\mathbf{x})$ gives, through straightforward calculations and the use of Lemma 3 and Equation (15),

$$
B_{j,\ell}^{(n)}(\mathbf{x}) = \frac{||K||_2^2}{nh_n^d}\frac{\bar{F}^{(1)}\left(y_{n,j}\vee y_{n,\ell}|\mathbf{x}\right)}{g(\mathbf{x})\bar{F}^{(1)}\left(y_{n,j}|\mathbf{x}\right)\bar{F}^{(1)}\left(y_{n,\ell}|\mathbf{x}\right)}(1+o(1)).
$$

The regular variation property of $\bar{F}^{(1)}$ gives $B_{j,\ell}^{(n)}(\mathbf{x}) = \frac{||K||_2^2}{g(\mathbf{x})}(\tau_j \vee \tau_\ell)^{-1}\epsilon_n^2(1+o(1))$. It remains to calculate $B_{j,J+1}^{(n)}(\mathbf{x})$. Using Equation (15), with $Q(\cdot) = K(\cdot)^2/||K||_2^2$ a kernel satisfying $(\mathcal{K})$, this term equals

$$
\frac{\frac{1}{nh_n^{2d}}||K||_2^2\mathbb{E}\left[|Y-z_n|^{p-1}Q\left(\frac{\mathbf{x}-\mathbf{X}}{h_n}\right)\mathbb{1}_{\{Y>z_n\vee y_{n,j}\}}\right]}{g(\mathbf{x})^2 B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)z_n^{p-1}\bar{F}^{(1)}(y_{n,j}|\mathbf{x})\bar{F}^{(1)}(z_n|\mathbf{x})/\gamma(\mathbf{x})(1+o(1))}
$$

$$
-\frac{\frac{1}{n}\mathbb{E}\left[\frac{1}{h_n^d}K\left(\frac{\mathbf{x}-\mathbf{X}}{h_n}\right)\mathbb{1}_{\{Y>y_{n,j}\}}\right]\mathbb{E}\left[|Y-z_n|^{p-1}\frac{1}{h_n^d}K\left(\frac{\mathbf{x}-\mathbf{X}}{h_n}\right)\mathbb{1}_{\{Y>z_n\}}\right]}{g(\mathbf{x})^2 B\left(p,\gamma(\mathbf{x})^{-1}-p+1\right)z_n^{p-1}\bar{F}^{(1)}(y_{n,j}|\mathbf{x})\bar{F}^{(1)}(z_n|\mathbf{x})/\gamma(\mathbf{x})(1+o(1))}.
$$

Clearly, as a direct consequence of Lemma 3, the first term dominates. Remark that $z_n \vee y_{n,j} = (1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta)z_n(1+o(1))$ and combine Assumption $(\mathcal{K})$, the results of Lemma 4 (with $\lambda = (1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta)$), and the regular variation property of $\varphi^{(k)}(\cdot)$ (see Equation (15)) to find that the numerator of this first term is asymptotically equivalent to

$$\frac{||K||_2^2}{nh_n^d} g(\mathbf{x}) z_n^{p-1} \bar{F}^{(1)}(z_n|\mathbf{x}) \left[(p-1) IB \left((1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta)^{-1}, \gamma(\mathbf{x})^{-1} - p + 1, p - 1 \right) \right.$$
$$\left. + ((1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta) - 1)^{p-1} \left(1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta \right)^{-1/\gamma(\mathbf{x})} \right].$$

And finally $B_{j,J+1}^{(n)}(\mathbf{x})$ is asymptotically equivalent to

$$\frac{\tau_j^{-1} \gamma(\mathbf{x}) \frac{||K||_2^2}{g(\mathbf{x})} \epsilon_n^2}{B \left(p, \gamma(\mathbf{x})^{-1} - p + 1 \right)} \left[(p-1) IB \left((1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta)^{-1}, \gamma(\mathbf{x})^{-1} - p + 1, p - 1 \right) \right.$$
$$\left. + ((1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta) - 1)^{p-1} \left(1 \vee \tau_j^{-\gamma(\mathbf{x})}/\theta \right)^{-1/\gamma(\mathbf{x})} \right].$$

Therefore, $\mathbb{V}\mathrm{ar}[Z_n] = ||K||_2^2 \boldsymbol{\beta}^\top \boldsymbol{\Lambda}(\mathbf{x}) \boldsymbol{\beta}/g(\mathbf{x})(1 + o(1))$, where $\boldsymbol{\Lambda}(\mathbf{x})$ is given in Equation (16). It remains to prove the asymptotic normality of Z_n. For that purpose, we denote $Z_n = \sum_{i=1}^n Z_{i,n}$, where

$$Z_{i,n} = \frac{\epsilon_n^{-1}}{nh_n^d} \sum_{j=1}^J \beta_j \frac{K \left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n} \right) \mathbb{1}_{\{Y_i > y_{n,j}\}} - \mathbb{E} \left[K \left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n} \right) \mathbb{1}_{\{Y_i > y_{n,j}\}} \right]}{\varphi^{(0)}(y_{n,j}|\mathbf{x})}$$
$$+ \frac{\epsilon_n^{-1}}{nh_n^d} \beta_{J+1} \frac{|Y_i - z_n|^{p-1} K \left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n} \right) \mathbb{1}_{\{Y_i > z_n\}} - \mathbb{E} \left[|Y_i - z_n|^{p-1} K \left(\frac{\mathbf{x}-\mathbf{X}_i}{h_n} \right) \mathbb{1}_{\{Y_i > z_n\}} \right]}{\varphi^{(p-1)}(z_n|\mathbf{x})}.$$

Taking $\delta > 0$ sufficiently small and arguing as in the closing stages of the proof of Lemma 6 in [17], we find that $n\mathbb{E} \left[|Z_{1,n}|^{2+\delta} \right] = O \left(\epsilon_n^\delta \right) = o(1)$. Applying the classical Lyapunov central limit theorem concludes the proof. $\qquad\square$

Proposition 1 Assume that $(\mathcal{K})$, $(\mathcal{L})$ and $C_2 (\gamma(\mathbf{x}), \rho(\mathbf{x}), A(.|\mathbf{x}))$ hold. Let $y_n \to \infty$, $h_n \to 0$ and $z_n = \theta y_n (1 + o(1))$, where $\theta > 0$. Assume further that $\epsilon_n^{-2} = nh_n^d \bar{F}^{(1)}(y_n|\mathbf{x}) \to \infty$, $nh_n^{d+2} \bar{F}^{(1)}(y_n|\mathbf{x}) \to 0$, $\omega_{h_n}^{(3)}(y_n|\mathbf{x}) \to 0$ and there exists $\delta \in (0,1)$ such that $\epsilon_n^{-1} \omega_{h_n}^{(1)}((1-\delta)(\theta \wedge 1)y_n|\mathbf{x}) \log(y_n) \to 0$. If, for all $j \in \{1, \ldots, J\}$, the $y_{n,j} = \tau_j^{-\gamma(\mathbf{x})} y_n (1 + o(1))$ with $0 < \tau_1 < \tau_2 < \ldots < \tau_J \leq 1$ are such that $\epsilon_n^{-1} \omega_{h_n}^{(2)}((1+\delta)(\theta \vee \tau_1^{-\gamma(\mathbf{x})})y_n|\mathbf{x}) \to 0$, then, for all $p \in (1, \gamma(\mathbf{x})^{-1}/2 + 1)$, one has

$$\epsilon_n^{-1} \left\{ \left(\frac{\hat{\bar{F}}_n^{(1)}(y_{n,j}|\mathbf{x})}{\bar{F}^{(1)}(y_{n,j}|\mathbf{x})} - 1 \right)_{1 \leq j \leq J}, \left(\frac{\hat{\bar{F}}_n^{(p)}(z_n|\mathbf{x})}{\bar{F}^{(p)}(z_n|\mathbf{x})} - 1 \right) \right\} \xrightarrow{d} \mathcal{N} \left(\mathbf{0}_{J+1}, \frac{||K||_2^2}{g(\mathbf{x})} \boldsymbol{\Lambda}(\mathbf{x}) \right),$$

where $\boldsymbol{\Lambda}(\mathbf{x})$ is given in Equation (16).

Proof Notice that

$$\frac{\hat{\bar{F}}_n^{(p)}(u_n|\mathbf{x})}{\bar{F}^{(p)}(u_n|\mathbf{x})} - 1 = \left(\frac{\hat{\varphi}_n^{(p-1)}(u_n|\mathbf{x})}{\varphi^{(p-1)}(u_n|\mathbf{x})} - 1 \right) \frac{m^{(p-1)}(u_n|\mathbf{x})}{\hat{m}_n^{(p-1)}(u_n|\mathbf{x})} + \left(\frac{m^{(p-1)}(u_n|\mathbf{x})}{\hat{m}_n^{(p-1)}(u_n|\mathbf{x})} - 1 \right).$$

Lemma 2 and the Chebyshev inequality ensure that for all $p \in (1, \gamma(\mathbf{x})^{-1}/2 + 1)$ and $u_n \in \{y_{n,1}, \ldots, y_{n,J}, z_n\}$, $\hat{m}_n^{(p-1)}(u_n|\mathbf{x})/m^{(p-1)}(u_n|\mathbf{x}) - 1 = O_{\mathbb{P}}(1/\sqrt{nh_n^d})$, so that

$$
\epsilon_n^{-1}\left(\frac{\hat{\bar{F}}_n^{(p)}(u_n|\mathbf{x})}{\bar{F}^{(p)}(u_n|\mathbf{x})} - 1\right) = \epsilon_n^{-1}\left(\frac{\hat{\varphi}_n^{(p-1)}(u_n|\mathbf{x})}{\varphi^{(p-1)}(u_n|\mathbf{x})} - 1\right) + o_{\mathbb{P}}(1).
$$

Applying Lemma 6 concludes the proof. $\qquad\square$

6.2 Proofs of main results

Proof of Theorem 1 Let us denote $\mathbf{t} = (t_1, \ldots, t_J, t_{J+1})$ and focus on the probability

$$
\Phi_n(\mathbf{t}) = \mathbb{P}\left(\bigcap_{j=1}^{J}\left\{\sigma_n^{-1}\left(\frac{\hat{q}_n^{(1)}(\alpha_{n,j}|\mathbf{x})}{q^{(1)}(\alpha_{n,j}|\mathbf{x})} - 1\right) \leq t_j\right\} \bigcap \left\{\sigma_n^{-1}\left(\frac{\hat{q}_n^{(p)}(a_n|\mathbf{x})}{q^{(p)}(a_n|\mathbf{x})} - 1\right) \leq t_{J+1}\right\}\right).
$$

Set $y_n = q^{(1)}(\alpha_n|\mathbf{x})$, $y_{n,j} = q^{(1)}(\alpha_{n,j}|\mathbf{x})\left(1 + \sigma_n t_j\right)$ and $z_n = q^{(p)}(a_n|\mathbf{x})\left(1 + \sigma_n t_{J+1}\right)$. The technique of proof of Proposition 1 in [16] yields

$$
\begin{aligned}
\Phi_n(\mathbf{t}) = \mathbb{P}\Bigg(&\bigcap_{j=1}^{J}\left\{\sigma_n^{-1}\left(\frac{\hat{\bar{F}}_n^{(1)}\left(y_{n,j}|\mathbf{x}\right)}{\bar{F}^{(1)}\left(y_{n,j}|\mathbf{x}\right)} - 1\right) \leq \sigma_n^{-1}\left(\frac{\bar{F}^{(1)}\left(q^{(1)}(\alpha_{n,j}|\mathbf{x})|\mathbf{x}\right)}{\bar{F}^{(1)}\left(y_{n,j}|\mathbf{x}\right)} - 1\right)\right\} \\
&\bigcap \left\{\sigma_n^{-1}\left(\frac{\hat{\bar{F}}_n^{(p)}\left(z_n|\mathbf{x}\right)}{\bar{F}^{(p)}\left(z_n|\mathbf{x}\right)} - 1\right) \leq \sigma_n^{-1}\left(\frac{\bar{F}^{(p)}\left(q^{(p)}(a_n|\mathbf{x})|\mathbf{x}\right)}{\bar{F}^{(p)}\left(z_n|\mathbf{x}\right)} - 1\right)\right\}\Bigg).
\end{aligned}
$$

Second-order regular variation arguments similar to those of the proof of Proposition 1 in [16] give, for all $j \in \{1, \ldots, J\}$,

$$
\sigma_n^{-1}\left(\frac{\bar{F}^{(1)}\left(q^{(1)}(\alpha_{n,j}|\mathbf{x})|\mathbf{x}\right)}{\bar{F}^{(1)}\left(y_{n,j}|\mathbf{x}\right)} - 1\right) = \frac{t_j}{\gamma(\mathbf{x})}(1 + o(1))
$$

and similarly

$$
\sigma_n^{-1}\left(\frac{\bar{F}^{(p)}\left(q^{(p)}(a_n|\mathbf{x})|\mathbf{x}\right)}{\bar{F}^{(p)}\left(z_n|\mathbf{x}\right)} - 1\right) = \frac{t_{J+1}}{\gamma(\mathbf{x})}(1 + o(1)).
$$

Finally, notice that $y_{n,j} = \tau_j^{-\gamma(\mathbf{x})}y_n(1+o(1))$ and $z_n = \theta y_n(1+o(1))$ (see (9)). Moreover, for n large enough, $\omega_{h_n}^{(1)}(y_{n,j}|\mathbf{x}) \leq \omega_{h_n}^{(1)}\left((1-\delta)q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}\right)$ and $\omega_{h_n}^{(1)}(z_n|\mathbf{x}) \leq \omega_{h_n}^{(1)}\left((1-\delta)\theta q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}\right)$. Similarly, $\omega_{h_n}^{(2)}(y_{n,j}|\mathbf{x}) \leq \omega_{h_n}^{(2)}\left((1+\delta)\tau_1^{-\gamma(\mathbf{x})}q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}\right)$ and $\omega_{h_n}^{(2)}(z_n|\mathbf{x}) \leq \omega_{h_n}^{(2)}\left((1+\delta)\theta q^{(1)}(\alpha_n|\mathbf{x})|\mathbf{x}\right)$. Conclude using Proposition 1. $\quad\square$

Proof of Theorem 2 We recall $\sigma_n^{-2} = n h_n^d (1 - \alpha_n)$. Write

$$\frac{\sigma_n^{-1}}{\log\left(\frac{1-\alpha_n}{1-\beta_n}\right)} \log\left(\frac{\tilde{q}_{n,\alpha_n}^{(p)}(\beta_n|\mathbf{x})}{q^{(p)}(\beta_n|\mathbf{x})}\right) = \sigma_n^{-1}(\hat{\gamma}_{\alpha_n}(\mathbf{x}) - \gamma(\mathbf{x})) + \frac{\sigma_n^{-1}}{\log\left(\frac{1-\alpha_n}{1-\beta_n}\right)} \log\left(\frac{\hat{q}_n^{(p)}(\alpha_n|\mathbf{x})}{q^{(p)}(\alpha_n|\mathbf{x})}\right)$$
$$+ \frac{\sigma_n^{-1}}{\log\left(\frac{1-\alpha_n}{1-\beta_n}\right)} \log\left(\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\gamma(\mathbf{x})} \frac{q^{(p)}(\alpha_n|\mathbf{x})}{q^{(p)}(\beta_n|\mathbf{x})}\right).$$

The first term converges in distribution to Γ. The second one converges to 0 in probability, by Theorem 1. To control the third one, write

$$\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\gamma(\mathbf{x})} \frac{q^{(p)}(\alpha_n|\mathbf{x})}{q^{(p)}(\beta_n|\mathbf{x})} = \left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\gamma(\mathbf{x})} \frac{q^{(1)}(\alpha_n|\mathbf{x})}{q^{(1)}(\beta_n|\mathbf{x})} \frac{q^{(p)}(\alpha_n|\mathbf{x})}{q^{(1)}(\alpha_n|\mathbf{x})} \frac{q^{(1)}(\beta_n|\mathbf{x})}{q^{(p)}(\beta_n|\mathbf{x})}.$$

In view of Theorem 4.3.8 in [11] and its proof, $((1 - \alpha_n)/(1 - \beta_n))^{\gamma(\mathbf{x})} q^{(1)}(\alpha_n|\mathbf{x}) = q^{(1)}(\beta_n|\mathbf{x}) \left(1 + O\left(A\left((1 - \alpha_n)^{-1}|\mathbf{x}\right)\right)\right) = q^{(1)}(\beta_n|\mathbf{x}) (1 + O(\sigma_n))$. By Lemma 5 then:

$$\left(\frac{1-\alpha_n}{1-\beta_n}\right)^{\gamma(\mathbf{x})} \frac{q^{(p)}(\alpha_n|\mathbf{x})}{q^{(p)}(\beta_n|\mathbf{x})} = 1 + O(\sigma_n).$$

The third term therefore converges to 0. Conclude using Slutsky's lemma and the delta-method. $\square$

Proof of Theorem 3 This proof is similar to those of Theorem 4 in [17] (where $p = 2$) and Theorem 1 in [16] (an unconditional version) and is thus left to the reader.

References

1. Artzner, P., Delbaen, F., Eber, J., and Heath, D. Coherent measures of risk. *Mathematical Finance 9*, 3 (1999), 203–228.
2. Bellini, F., Klar, B., Muller, A., and Gianin, E. R. Generalized quantiles as risk measures. *Insurance: Mathematics and Economics 54* (2014), 41–48.
3. Cai, J., and Weng, C. Optimal reinsurance with expectile. *Scandinavian Actuarial Journal 2016*, 7 (2016), 624–645.
4. Chen, Z. Conditional L_p-quantiles and their application to the testing of symmetry in non-parametric regression. *Statistics & Probability Letters 29*, 2 (1996), 107–115.
5. Daouia, A., Gardes, L., and Girard, S. On kernel smoothing for extremal quantile regression. *Bernoulli 19*, 5B (2013), 2557–2589.
6. Daouia, A., Gardes, L., Girard, S., and Lekina, A. Kernel estimators of extreme level curves. *TEST 20*, 2 (2011), 311–333.
7. Daouia, A., Girard, S., and Stupfler, G. Estimation of tail risk based on extreme expectiles. *Journal of the Royal Statistical Society: Series B 80*, 2 (2018), 263–292.
8. Daouia, A., Girard, S., and Stupfler, G. Extreme M-quantiles as risk measures: from L^1 to L^p optimization. *Bernoulli 25*, 1 (2019), 264–309.
9. Daouia, A., Girard, S., and Stupfler, G. ExpectHill estimation, extreme risk and heavy tails. *Journal of Econometrics*, to appear, 2020.

10. Daouia, A., Girard, S., and Stupfler, G. Tail expectile process and risk assessment. *Bernoulli 26*, 1 (2020), 531–556.

11. de Haan, L., and Ferreira, A. *Extreme Value Theory: An Introduction*. Springer-Verlag New York, 2006.

12. El Methni, J., Gardes, L., and Girard, S. Non-parametric estimation of extreme risk measures from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics 41*, 4 (2014), 988–1012.

13. El Methni, J., and Stupfler, G. Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions. *Statistica Sinica 27*, 2 (2017), 907–930.

14. Embrechts, P., Kluppelberg, C., and Mikosch, T. *Modelling Extremal Events*. Springer, 1997.

15. Gardes, L., Girard, S., and Stupfler, G. Beyond tail median and conditional tail expectation: extreme risk estimation using tail L^p−optimisation. *Scandinavian Journal of Statistics* (2019), to appear.

16. Girard, S., Stupfler, G., and Usseglio-Carleve, A. An L^p-quantile methodology for estimating extreme expectiles, preprint. https://hal.inria.fr/hal-02311609v3/document, 2019.

17. Girard, S., Stupfler, G., and Usseglio-Carleve, A. Nonparametric extreme conditional expectile estimation, preprint. https://hal.archives-ouvertes.fr/hal-02114255, 2019.

18. Jones, M. Expectiles and M-quantiles are quantiles. *Statistics & Probability Letters 20* (1994), 149–153.

19. Koenker, R., and Bassett, G. J. Regression quantiles. *Econometrica 46*, 1 (1978), 33–50.

20. Newey, W., and Powell, J. Asymmetric least squares estimation and testing. *Econometrica 55*, 4 (1987), 819–847.

21. Ohlsson, E., and Johansson, B. *Non-Life Insurance Pricing with Generalized Linear Models*. Springer, 2010.

22. Usseglio-Carleve, A. Estimation of conditional extreme risk measures from heavy-tailed elliptical random vectors. *Electronic Journal of Statistics 12*, 2 (2018), 4057–4093.

23. Weissman, I. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association 73*, 364 (1978), 812–815.